面向南海问题的智库知识库构建研究 *

■ 丁晟春 史金晶
南京理工大学经济管理学院 南京 210094

摘要：［目的 / 意义］面向南海问题的智库能够更好地服务国家南海战略，是新时期我国南海问题研究的重点领域。构建南海问题智库知识库，可以实现对南海问题知识的有效组织与利用，为智库产生高质量产品提供知识保障。［方法 / 过程］本文引入本体理论提出了南海问题知识库的构建方法。在信息源分析和需求分析的基础上，设计了知识库的总体结构，构建了南海问题本体，并以“中菲黄岩岛对峙事件”为例实现了知识库的构建。［结果 / 结论］实证分析结构表明，本文所构建的南海问题知识库能够较好地组织存储相关知识，为南海问题智库的研究提供有益参考，所提出的基于本体的知识库构建方法可以将知识有效地整合，是在专业智库知识库构建领域的一次有益尝试。

关键词：南海问题 智库 知识库 本体
分类号：C931.6
DOI: 10.19318/j.cnki.issn.2096-1634.2016.02.10

在当今世界，智库已经成为一个现代国家决策链条上不可缺少的重要一环，是影响政府决策制定的重要力量。有学者将其称为是继立法、行政、司法和媒体之后的“第五权利中心”[1]。2013 年 11 月，党的十八届三中全会通过《中共中央关于全面深化改革若干重大问题的决定》，正式提出要“加强中国特色新型智库建设，建立健全决策咨询制度”，这是推进我国国家治理体系现代化的重要举措。在这个契机下，我们应该就我国目前战略发展的目标与挑战，建立一批新型的专业化智库，为国家的科学决策提供有力的智力支撑。

“南海问题”主要是指中国南海诸岛主权及其附属海域权益归属问题，是中国与南海周边国家关系中一个复杂而重要的议题。由于南海是太平洋上重要的战略军事要地，南海问题受到了国际上多方力量的关注和影响，国家间的博弈使得南海局势的稳定和问题的最终解决变得更加艰难。因此如何制定出南海问题的最佳解决方案成为我国政府面临的一项重要任务。建立面向南海问题的智库，对南海问题进行深入研究，可以更好地服务我国南海战略。智库高质量产品的“质变”依赖于知识资源的组织、利用和创造[2]。智库作为知识型的组织机构，有效的知识组织和知识管理是智库研究的基础保障。如何对搜集到的数量庞大、资源异构的信息资源进行科学地组织管理是一个需要探讨的问题。TTCSP（Think Tanks and Civil Societies Program）全球性报告中特别强调，迅速变化的信息环境和竞争激烈的行业环境要求智库构建完善的知识库以应对挑战。因此本文将知识工程领域的知识库理论应用到智库中来，面向南海问题构建知识库。

1 相关工作

知识库（Knowledge Base）是关于某一领域

* 本文系国家社会科学基金项目“基于社会网络分析的网络舆情主题发现研究”(编号：15BTQ063)，国家社会科学基金重点项目“大数据环境下社会舆情与决策支持方法体系研究”(编号：14AZD084) 研究成果之一，获得中央高校基本科研业务费专项资金资助 (编号：30916011330)。
作者简介：丁晟春 (ORCID: 0000-0002-4269-021X)，副教授，E-mail: todingding@163.com; 史金晶 (ORCID: 0000-0002-7547-2884)，硕士研究生，E-mail: shijinjing118@163.com。

各种知识的集合，用于实现知识的存储、管理和重用 [3]。就其内涵而言，知识库是一种基于数据库技术，针对特定领域问题求解的需要进行知识抽取，并且根据一定的知识体系实现知识分类和组织后，针对知识特点和需求存储在计算机中的相互关联的有序化的知识集合，是一种专业化的面向用户需求的知识服务系统 [4]。许鑫参照知识库的定义，将智库知识库表述为，通过积累来自各种渠道的和保留自身研究产生的各类信息资源，支持研究人员进行产品开发和咨询服务的平台工具 [5]。

国外对于智库资源建设的研究起步较早，研究比较成熟，许多智库已经建立了具有一定规模的知识库。美国兰德公司根据不同的研究领域设立了专题知识库，包括全球恐怖事件数据集、健康和生育档案等。瑞士的斯德哥尔摩国际和平研究所建立了许多大型数据库，包括国际关系与安全数据库、军费开支数据库、武器转让数据库等。美国国际经济研究局建立了多种类型的数据库，包括宏观数据、产业数据、医疗数据、专利与科学论文数据等，并且将这些数据在网站公布，供研究人员引用和参考 [6]。

现阶段国内研究大多集中在智库的社会职能方面，针对智库知识库和智库知识资源方面的研究关注较少。殷志华等论述了建立智库知识管理系统的必要性，认为智库研究人员需要获取和共享分类清晰的知识文献，对历史情况进行纵向比较分析以及与国际情况横向比较分析。获取信息资源是政策研究的前提条件，对基础信息资源进行系统管理是政策研究的必要条件 [7]。王世伟认为应该建立专题特色数据库进行知识积累，为智库的前端服务提供基础 [6]。许鑫通过对国外智库知识库案例进行分析研究，借鉴机构知识库的经验，提出了一种知识库构建通用流程，但是缺少实例验证 [5]。

知识库的构建研究是知识工程领域内非常重要的问题，国内外相关学者对知识库构建的理论与技术的研究近年来不断深化发展，取得了一定的成果并得到了有效地应用。目前主要有四种常规的构建方法：（1）基于聚类方法的知识库构建方法；（2）基于可拓知识表示的知识库构建方法；（3）基于神经网络的知识库构建方法；（4）基于本体的知识库构建方法。本体是一种共享的概念定义集，能在语义和知识层次上描述概念及概念间的关系，消除领域概念间的混乱与歧义，最大程度地实现知识的共享与重用。国内已有很多领域引入了基于本体的知识库，如产品设计、军事等领域。饶国政认为理想的知识库是应该建立在本体的基础之上的，将本体用于知识库建设有许多突出的优点。本体可以为知识库的构建提供一个基本的结构，既可以描述静态实体，也可以描述事物或概念的运动和变化。在逻辑建模中，本体的使用可以帮助人们清楚地理解特定领域的相关元素、关系和概念，让知识表达更加准确便捷，从而有助于更好地管理和维护领域知识库 [8]。

本文将构建南海问题本体，基于本体构建南海问题知识库，将南海问题知识进行组织，更好地服务于南海智库的研究。

2 南海问题知识库设计

构建南海问题知识库首先应对信息的来源及信息的要素进行总结归纳。其次针对南海问题研究的目的，分析目标用户对知识库的需求，明确知识库中需要存储的知识以及知识类型。在需求分析的基础上对知识库进行总体结构设计，根据南海问题知识类型的不同，设计知识库的存储模式。最后，完成南海问题本体库和数据库的构建。

2.1 南海问题知识库信息源分析

本文将南海知识库的信息源限定在网络平台上。根据网络信息的不同形式，本文将南海问题相关信息的获取来源总结归纳为网站和社交媒体两方面。

（1）网站是因特网上发布资讯和提供服务的主要平台，有关南海问题的网络信息主要来源于机构网站和新闻类网站。① 机构网站是相关机构在网络上发布信息的平台，可以分为政府机构网站和研究机构网站。政府机构网站是官方权威信息发布的主要平台，在这类网站上不仅可以获取南海问题相关的法律法规、最新的国家政策，还可以了解到官方对南海问题重大事件的明确态度。② 新闻类网站是综合性网络媒体，是以新闻为主的大型网上信息交互平台，具有报道完整、要素齐备等特征。新闻类网站会以专题的形式对重大

突发事件进行全方位的报道，为有关南海问题的信息收集提供便利。

（2）社交媒体是 web2.0 时代兴起的新的信息传播途径，主要包括论坛、博客、微博等。近年来微博的用户数量激增，影响力急剧扩大，逐渐成为网络信息传播最重要的载体。微博的传播速度非常快，突发事件一经曝光，短时间内就可以迅速扩散，形成全网大讨论的态势。例如“中菲黄岩岛对峙”事件发生后，微博实时更新事件的进展，网民就此事件展开了激烈的讨论。微博具有评论、转发、点赞等功能，民众可对相关事件发表观点意见，这些舆情信息是南海问题研究的重要内容。

2.2 南海问题知识库需求分析

南海问题知识库的目标用户主要是南海问题的研究人员，他们在分析研究过程中需要了解或参考大量的南海问题相关知识。这些底层的信息资源直接影响南海问题分析结果的科学性与客观性，所以对相关知识的完备性与准确性就提出了较高的要求。目前与南海问题相关的知识来源非常广泛、内容多样、数量众多，并且不断地变化更新。本文根据知识类别的不同，将南海问题研究人员的知识需求总结为以下三个方面：

（1）南海问题基本信息

南海问题涉及的岛屿及海域的概况：南海是指中国确立九段线内群岛及其附属海域。基本信息主要包括南海诸岛及海域的行政划分、岛屿的地理位置、岛屿名称、目前实际控制状态等。

南海问题的参与主体：南海问题主要涉及南海周边六国，包括中国、菲律宾、文莱、马来西亚、越南和印度尼西亚，在南海问题研究中被称为南海声索国。近来南海问题涉及的主体开始多元化，美国、日本等大国以“自由航行”等借口介入到南海主权争端中。此外，东南亚国家联盟是南海局势稳定的重要影响力量，也成为需要研究的主体之一。

（2）南海争端事件信息

由于历史遗留等原因，中国与周边国家在南海主权问题上的争端时有发生。这些争端事件是南海问题研究的重点，研究人员需要了解事件的相关信息及发展态势。包括事件的基本信息，如发生的时间、地点、人物、原因、经过、内容等。同时事件的传播信息也是分析的重要内容，包括事件衍生出的舆情信息、事件关注者的基本信息等。

（3）南海问题文献信息

对于南海问题研究开始于 20 世纪 70 年代，从多学科的视角进行深入研究，可总结为历史地理研究领域、法律与国际因素研究领域、中国视角对策研究领域、非传统安全研究领域等。目前国内已经涌现了一大批该领域优秀的专家学者，建立了一批以南海问题为研究方向的学术机构 [9]。国外研究机构对南海问题也相当关注，主要研究集中在南海问题的根源、解决方案、影响因素和前景等几方面 [9]。这些学术研究成果是南海智库重要的知识支撑，因此在南海问题知识库中应该存储有关的文献信息。

2.3 南海问题知识库结构设计

南海问题知识库的总体结构包括本体库和数据库两个部分，总体框架如图 1 所示。本体库中存储南海地理信息、南海争端事件信息、南海问题中涉及到的国家与关注方的信息，这些信息通过南海争端事件本体建立语义关联。

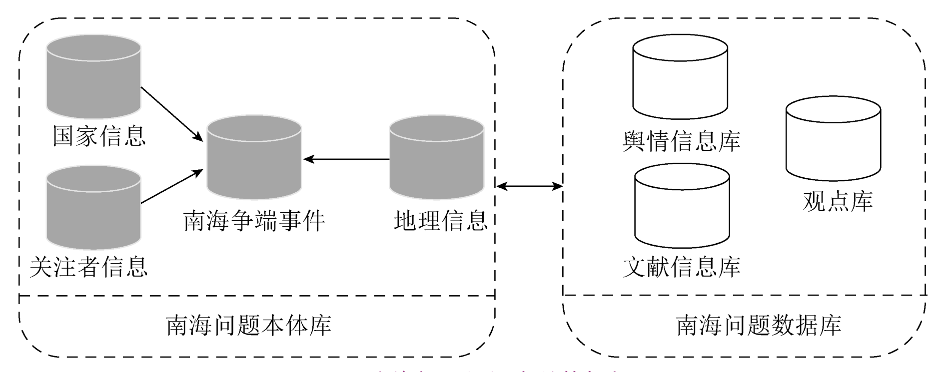


图 1 南海问题知识库总体框架

Figure 1 The framework of the South China Sea issue knowledge base

南海问题数据库包括舆情信息库、文献信息库和观点库三个子库，舆情信息库负责存储收集舆情信息，如网页、博客、微博和网民评论等，文献信息库存储书籍、期刊等相关文献资源的著录信息，观点库存储通过相关舆情分析技术处理得到的关注者的观点立场信息。

3 南海问题本体库构建

本文根据南海问题智库研究的实际情况和需要，以斯坦福大学的七步法为主体提出了南海问题本体的构建方法，其步骤如下：① 确定南海问题本体的应用领域，主要用于南海问题的研究，为南海问题智库提供知识支持。② 考虑现有事件本体的复用性。本文参考 ABC 共享本体的语义规范 [10] 来指导南海问题本体的构建。③ 列出领域中重要概念和术语。列举出南海问题研究领域的重要概念和术语。④ 定义类与类的层次。采用自顶向下的方法定义南海问题本体中类的层次结构。⑤ 根据类的特征定义数据属性和对象属性，规定属性的约束条件。⑥ 创建实例。在完成上述步骤后为定义的每个类创建实例，并且添加该实例具体的属性值。⑦ 本体的检验及评价。检验构建的本体中是否存在错误，是否能满足本体构建最初的目标。

3.1 本体概念和类的提取

依据上文所确定的研究领域与范畴，在知识分类的基础上，参考文献资料，对核心概念进行提取。在此基础上，对这些概念进行评估，按照一定的逻辑规则分组，确定从南海争端类、国家类、岛屿类、关注者类这四个方面构建南海问题本体，如图 2 所示。然后对每一组概念自顶向下进行分层。

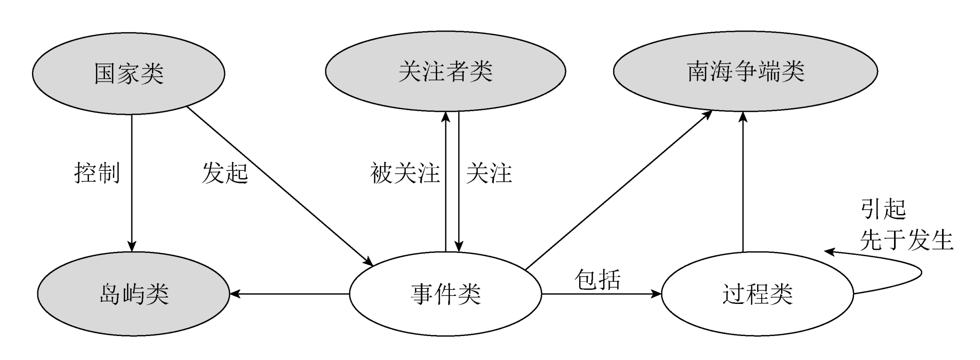


图 2 南海问题本体类
Figure 2 The classes of the South China Sea issue ontology

（1）南海争端类，是指南海争端事件，将其分为事件类和过程类。一个事件有一个或多个过程。事件过程类表示一个动态事件的概念，表示事件在某个时间区域上的划分，刻画突发事件在时间演进中的状态。

（2）国家类，是指南海问题所涉及到的国家。本文以国家在南海问题上的角色为分类标准，将国家类分为利益主体类和一般国家类这两个二级子类。在利益主体类下分为南海声索国类、外部利益国类、联盟类这三个三级子类。具体的分类体系如图 3 所示。

（3）关注者类，指对南海问题进行报道研究的个人或机构。本文将关注者类分为个人类、研究机构类、媒体类、政府类这四个二级子类。分类体系见图 4。

（4）岛屿类，是对南海诸岛的概括。本文根据已有的地理区域划分，分为东沙群岛、西沙群岛、南沙群岛、中沙群岛四个二级子类。

3.2 数据属性的定义

数据属性是表示类或概念与值的关系，在前期工作的基础上，本文对所涉及概念的数据属性进行定义，定义其数据属性的约束，用来描述属性的值类型、允许值（领域和范围）、值基准、以及值的其他特性，表 1 为部分南海问题本体库的数据属性。

3.3 非等级关系的定义

本体的对象属性用来表示类之间的非等级关系，表2为部分南海问题本体库的非等级关系的说明。

丁晟春，史金晶．面向南海问题的智库知识库构建研究 [J]. 智库理论与实践，2016，1(2)：63-69.

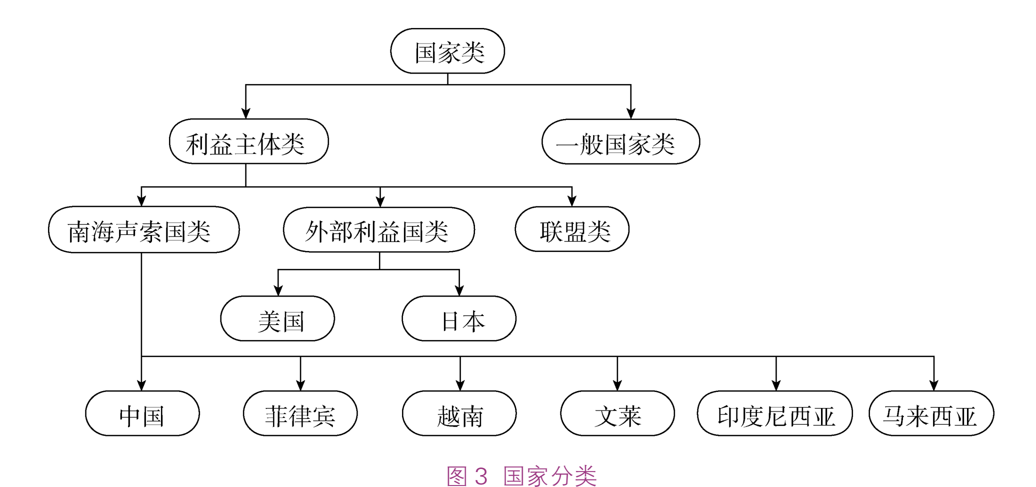


图 3　国家分类

Figure 3　The classification of country

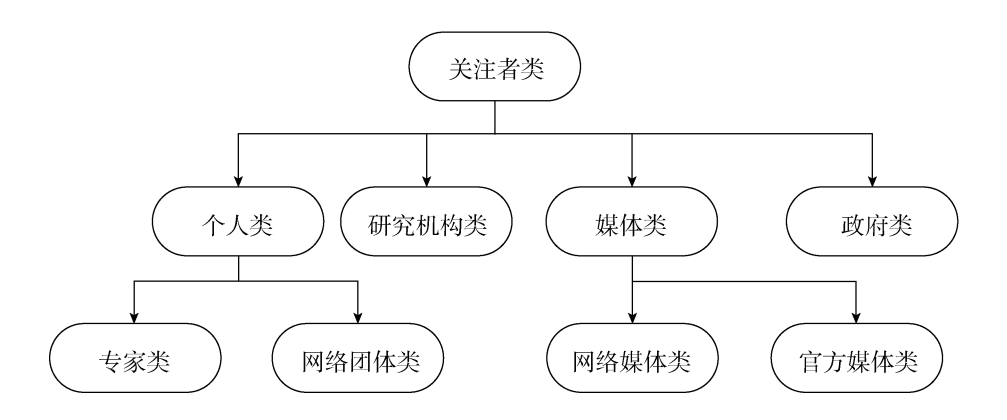


图 4　关注者分类

Figure 4　The classification of follower

表 1　部分本体数据属性

Table 1　The part of data properties

类名	数据属性
国家类	编号 (number)，全称（name），英文全称 (nameen)，别称（alias），首都 (capital)
岛屿类	编号 (number)，中文名称（name），英文名称 (nameen)，别名 (alias)，所属地区 (ownership)，地理位置 (location)
南海争端类	编号 (number)，名称（name），起始时间 (starttime)，结束时间 (endtime)，参与主体 (agent)，受影响客体 (object)，事件性质 (kind)，手段 (means)，内容 (contents)，来源（sources）
关注者类	编号 (number)，中文名称 (name)，英文名称 (nameen)，网址（url），创办时间 (foundation)，所属单位 (ownership)，主办单位 (organizer)，总部地点 (location)，经营范围 (range)，创始人 (builder)，负责人 (chairman)

3.4 实例的添加

根据构建的本体结构，本文以“中菲黄岩岛对峙”事件为例，使用 protégé4.3 对南海问题本体进行实例的添加，同时验证了本体的有效性。图 5 是南海争端类下事件类和过程类的实例截图，显示了“中菲黄岩岛对峙”事件实例、“中国海监船抵黄岩岛”过程实例下的所有属性值。

南海问题本体主要是对南海问题相关信息进行表示，在本体中体现为类、数据属性、对象属性与实例。本文将本体以 OWL 文本的格式保存在本地计算机中，采用这种存储方式进行本体的存储便于复制与备份，并且可以随时进行编辑和修改。

表 2 部分本体非等级关系

Table 2 The part of object properties

	关系名称	关系说明	所属类	关联类
互逆	包括 include	表示抽象定义的事件可以包含多个过程	事件类	过程类
	是 …… 的过程 is_Part_Of	表示多个过程完成了一件事件	过程类	事件类
互逆	关注了 comment	表示关注者关注了哪些南海争端	关注者类	南海争端类
	被关注 is_Commented_By	表示某一南海争端被哪些关注者关注了	南海争端类	关注者类
互逆	介入 affect	表示某国家利用国际影响力或者是为了追求私利而介入某事件	利益主体类	南海争端类
	被介入 affected_By	表示该事件影响某国	南海争端类	利益主体类
互逆	发起 raise	表示国家主动发起某一南海争端	国家类	南海争端类
	发起国 is_Raised_By	表示某一南海争端是由哪些国家发起的	南海争端类	国家类
	发生地点 locate	表示争端或侵占对象岛屿的情况	南海争端类	岛屿类
	继 …… 之后 occur_after	表示过程之间先后顺序发生的关系	过程类	过程类
	国籍 nationality	表示关注者的国籍	关注者类	国家类



图 5 本体实例

Figure 5 The individuals

4 结语

在当今大数据背景下，构建知识库来管理智库信息资源，对知识进行有效的搜集、组织、存储和管理，为智库决策提供坚实的知识基础，这是智库未来发展所必须解决的一个问题。面向特定领域研究的专业智库，由于研究范围的不同，对知识资源的需求侧重点有所区别，因此构建专业化的知识库很有必要。本文基于本体的方法构建南海问题知识库，提出了具体的构建流程，是在专业智库知识库构建领域的一次尝试，希望在后续的研究中加以改进和完善。

参考文献：

[1] 里奇 . 智库、公共政策和专家治策的政治学 [M]. 潘羽辉，等译 . 上海 : 上海社会科学元出版社 , 2010: 6

[2] 李纲 , 李阳 . 情报视角下的智库建设研究 [J]. 图书情报工作 , 2015(11): 36-41.

[3] 钱智勇 , 周建忠 , 贾捷 . 楚辞知识库构建与网站实现研究 [J]. 图书馆理论与实践 , 2010(10): 70-73.

[4] 王新昊 . 情报研究工作中知识地图和知识库的构建研究 [D]. 南京 : 南京理工大学 , 2005.

[5] 许鑫 , 吴珊燕 . 智库知识库的构建研究 [J]. 情报理论与实践 , 2014, 37: 68-72.

[6] 王世伟 . 试析情报工作在智库中的前端作用——以上海社会科学院信息研究所为例 [J]. 情报资料工作 , 2011, 32(2): 92-96.

[7] 殷志华 , 赵国辉 . 利用信息化手段建立智库研究成果知识管理系统浅析 [J]. 中国建设信息 , 2014, 17: 66-67.

[8] 饶国政 , 冯志勇 , 张丽 , 等 . 基于本体的物流行业知识库的研究 [J]. 计算机应用研究 , 2008, 25(9): 2734-2737.

[9] 曾勇 . 国外南海问题研究述评 [J]. 现代国际关系 , 2012(6): 56-64.

[10] Corda I., Bennett B., Dimitrova V., A Logical Model of an Event Ontologyfor Exploring Connections in Historical Domains[C]. Proceedings of the 10th International Semantic Web Conference (ISWC), Bonn, Germany, 2011.

作者贡献说明：

丁晟春：提出论文思想，确立论文结构，修改论文；

史金晶：构建知识库模型，撰写论文。

Research on the Construction of the South China Sea Issue Oriented Think Tank Knowledge Base

Ding Shengchun Shi Jinjing

School of Economics and Management, Nanjing University of Technology and Engineering, Nanjing 210094

Abstract: **[Purpose/significance]** The think tank of the South China Sea issue can contribute to the development of the national strategy. It is the focus of the South China Sea issue research in the new era. The construction of the South China Sea issue oriented think tank knowledge base is helpful for the organizing and using of knowledge. It can provide a high-quality product for the think tank. **[Method/process]** We introduced a method of constructing the ontology theory knowledge of the South China Sea issue. On the basis of the information source analysis and the demand analysis, we designed the overall structure of the knowledge base and the South China Sea issue ontology. Besides, we took the "Philippine Huangyan Island confrontation" as an example to realize the construction of the repository. **[Result/conclusion]** The empirical analysis of the structure shows that the South China Sea issue oriented knowledge base can better organize storage-related knowledge, and provide a useful reference for the study of the South China Sea issue think tank. The knowledge-based ontology construction method achieved good effect, and it is a useful attempt.

Keywords: the South China Sea issue think tank knowledge base ontology

收稿日期：2015-10-22 修回日期：2016-04-23 本文责任编辑：吕青